Genome wide Identification and Expression analysis of WRKY transcription factors in *Avena sativa* L. under different stress condition

Synopsis

Submitted in the partial fulfilment for the requirement of

DOCTOR OF PHILOSOPHY

In

BOTANY

Submitted By

Hunny Waswani

Under the Supervision of

Dr. Rajiv Ranjan

Associate Professor

Department of Botany, Faculty of Science

Dayalbagh Educational Institute (Deemed University)

Dayalbagh, Agra-282005

2022

Introduction

Globally, crops aresubjected to several abiotic stresses that adversely affect crop productivity. More than 50% of crop losses each year are caused by abiotic stress on plants (Shao *etal*. 2015). Climate change studies have shown that crop plants will have a greater chance of surviving in these adverse environmental stresses in future, since these conditions will have a serious impact on their growth and development (Pandey *etal*. 2017).

Various abiotic factors such as heat, salinity and drought can negatively affect plants. Heat and drought stress influence multiple aspects of plant growth and development. Both drought and heat negatively affect plant growth, which leads to leaves withering, fading, damage of cell membrane, reduced leaf size and number, decline in germination rate, shorter stem length, reduction in photosynthetic rate, membrane integrity alteration in the end affecting productivity of crop (Lamaoui*etal*. 2018; Wang *etal*. 2019).

Furthermore, salinity poses a major threat to agricultural productivity across the world. Crops can't adapt themselves to salinity stress, global yields for all glycophytic crops are reduced by 50–80% when moderate salinity exists (Panta*et al*. 2014, Swarajyalakshmi*et al.* 2003). Approximately 20% of the cultivated land contributes to one third of all food produced worldwide, but the increased salinity affects about 70% of the cropland. In a study conducted in 1994, it was observed that a drop in rainfall of 10% causes an average 4.2% decline in cereal yield (Aredo 1995). All water constraints, including droughtresult in losses of 15–30% in agricultural output. In boosting plant defence responses against biotic and abiotic stresses, a pivotal role is carried out by phytohormones such as Abscisic acid (ABA), Gibberellic Acid (GA), Salicylic Acid (SA), Methyl Jasmonate (MeJA).Additionally, phytohormone signalling during stress is governed by transcription factors that respond to stress(Nolan *etal*. 2017).

To cope up with these environmental stresses plants find a way to escape from these adverse conditions by following robust mechanism with integrated transcription factors, hormonal factors(Beddington *etal*. 2012).Target gene's expression level is guided by the regulatory elements known as Transcription Factor (TFs), These TFs are specified protein classes, which bind to the *cis*-acting elements which is specific sequence site in promoter region of target gene, these regulatory elements perform a vital role in growth and development of plant, and plant's response to various abiotic stress like drought, cold,

heat and salinity. Mainly plant genes are induced at the transcription level, their expression shows a spatio-temporal regulation which is a major component that how plant respond to various environmental stresses.

***Avena sativa* L.**

A grain crop or an annual grass belongs to the family Poaceae (Gramineae). The oat plant has fibrous roots that grow into hollow stems 2-4 feet high with rounded pale green leaves with a narrow, flat texture. World-wide, oats are the fifth most widely cultivated grains.It is estimated that the size of Genome is 12.3 GB. Majorly two species of oats are produced worldwide namely: *Avena sativa* L. and *Avena nuda* L. The most widely grown oat species in the world is *Avena sativa* L., followed by *Avena nuda* L., *Avena sativa* L. is widely grown in moist and cool regions of North America and Northern Europe, Within the European union Poland and Finland are the largest producers of oat. while *Avena nuda* L. is the most widely grown species in other continents. In India,Uttar Pradesh, Punjab, Haryana, West Bengal and in some parts of Madhya Pradesh are major oat growing states.

For millennia, the cultivation of oats (*Avena sativa* L.) has been considered a beneficial source of food. The positive attributes of oats are their abundant supply of carbohydrate, including starch and soluble fiber, as well as their high level of lipids, protein, and B vitamins. The soluble dietary fibres in oat seeds reduces serum cholesterol levels and may prevent coronary heart disease. Also, they have higher components for health promoting factors with varied concentrations of amino acids, crude fat, and soluble fat. Life cycle of oat plant is same as the other annual grasses like wheat, rye, and barley.

In contrast to soybeans, maize, or wheat crops, thecommercial production of this crop is profitable, and it is usually cultivated in areas that have several disadvantages, such as drought or high salinity(Bai *etal.* 2018). The species is also sensitive to changes in light temperature (Heuschele*etal.* 2019). On a molecular level, our understanding of the mechanisms of stress adaptation in oat is not yet explored fully, and the main reason is that the genomes of the most important crops, like wheat(Appels *etal.* 2018), maize(Schnable*etal.* 2009) and rice(Goff *etal*. 2002; Yu *etal.* 2002)have been sequenced first. Therefore, these species could be better investigated. Due to its health benefits, oat is becoming more and more popular these days. Recently (in 2020), the oat genome has been sequenced and assembled for the first time, a technical breakthrough that will revolutionize oat research.

To better understand genes involved in oat's abiotic stress response and examining how they're expressed under adverse environment conditions, will help us better understand how these genes function. Several studies have been conducted in the past few years to examine oat reference gene stability under various conditions. Despite this, no such studies have been reported that could provide extensive insight upon the role of transcription factor superfamily WRKY.In plants, WRKY transcription factors are important for signal transduction networks that regulate a wide array of plant processes. They also form part of the largest family of transcriptional regulators.

Structural Characteristics of WRKY TFs

One of the largest families of transcription factors found in plants is the WRKY family. They play an integral part in signalling networks that modulate a significant number of plant processes. Conserved domain of WRKY transcription factor superfamily is WRKYGQK, along with (C-C, H-H/C) a zinc finger motif (Eulgem*etal.* 2000). Based on domains of WRKY and zinc finger motif's arrangement, the classification of transcription factor WRKY is majorly done in three groups: At both N and C terminal along with C2H2 zinc finger motif (CX4-5C X22-23HXH) Group I protein sequences contain two WRKY domains, Group II proteins only have one WRKY domain which is followed by C2H2 zinc finger motif (CX4-5CX22-23HXH). Also Group II is further sub divided into 5 groups which is based on the characteristics of sequence and phylogenetic analysis namely IIa, IIb, IIc, IId and IIe. Likewise, II group, IIIgroup proteins sequences have WRKY domain however, it conserves a C2HC zinc-finger motif (CX7CX23HXC) instead of a C2H2 motif. (Bakshi and Oelmüller, 2014; Rushton *etal.*2015**)**

Review of Literature

Various sophisticated mechanisms developed by plants to cope up with climate extremes and particularly by involving Transcription Factors (TFs). This study demonstrates, a genome-wide screening for identifying and analyze the WRKY TFs in pearl millet (*Pennisetumglaucum*; PgWRKY), which is one of the most widely grown cereal crops in the semi-arid regions. A total number of 97 putative PgWRKY proteins were identified and classified into three major Groups (I-III) based on the presence of WRKY DNA binding domain and zinc-finger motif structures **(Chanwala*et al.* 2020).**

Essential roles in regulation of stress responses carries out by WRKY transcription factors. Th study recognized 79 WRKY genes in *Solanum tuberosum*(potato). The expression analysis indicated that

seven StWRKYs seemed to respond to abiotic stress and phytohormone like salicylic acid treatment. These 79 WRKY genes are candidates for abiotic stress signalling for future experiments**(Zhang *et al.*2017).**

Integral signalling participates among various stresses are carried out by Transcription factors (TFs) by constituting a network. Among the major TFs, WRKY proteins play pivotal roles in reprograming at transcription level and in regulation of associated responses of stress. In view of this, genome- wide identification and transcriptome of WRKYTF's family was performed in the plants that follows C4 pathways, *S. viridis* (*Sv*WRKY) and *Setariaitalica* (*Si*WRKY), respectively. The study identified 44 *Sv*WRKYand 105 *Si*WRKYproteins that were analyzed computationally for the further stress research work.**(Muthamilarasan*et al*. 2015).**

Alfalfa (*Medicago sativa* L.) is the most widely grown leguminous herb in the world. Its agricultural development has been constricted by different adverse environmental conditions, along with dehydration, elevation in salinity, and cold temperature. A total of 107 WRKY genes were identified in alfalfa and divided into 3 main groups also this study provides an important foundation for future cloning and functional studies of WRKY genes in alfalfa (**Mao *et al*. 2020**).

To date, WRKY genes in kiwifruit (*Actinidia spp*.) remain poorly understood. In this study, a total of 97 *Ac*WRKY genes have been identified in the kiwifruit genome, Furthermore, the analysis of selected *Ac*WRKY genes showed a variety of expression patterns in five different organs as well as during abiotic and biotic stresses. The genome-wide characterization and identification of *Actinidia spp*. WRKY transcription factors provide insight into the evolutionary history and is a useful resource for further functional analyses of kiwifruit (**Jing *et al*. 2018**).

This study performed, *In Silico*characterization and identification of transcription factor WRKY encoding genes in chickpea genome. For this purpose, a systematic genome-wide analysis was carried out to identify the non-redundant WRKY transcription factors in the chickpea genome. This study has computationally identified 70 WRKY-encoding non-redundant genes which were distributed on all the chickpea chromosomes except chromosome 8 (**Waqas *et al*. 2019**).

A wide variety of developmental and physiological processes are regulated by WRKY proteins, especially those related to survival under diverse biotic and abiotic conditions. In this study, 100 putative

*Ptr*WRKY genes encoded the proteins contained in the complete WRKY domain in*Populustrichocarpa*, these results indicate that woody plants can be improved in their stress resistance and stress tolerance(**Jiang *et al*. 2014**).

The past decade has seen significant progress towards functionally characterizing and identifying WRKY genes in various species, comprised knowledge is about the WRKY family in sorghum (*Sorghum bicolor* (L.) moench). In this study the comprehensive identification of 94 putative WRKY transcription factors (SbWRKYs) were done. The induction of SbWRKYs by different abiotic stresses hinted at their potential involvement in stress responses. qRT-PCR analysis revealed different expression patterns for SbWRKYs during drought stress (**Baillo*etal*. 2020**).

Cucumis sativus L. (Cucumber) is an economically important vegetable crop species. However, it is susceptible to various abiotic and biotic stresses, in this study, an analysis of the new assembly of the cucumber genome (v3.0) allowed the identification of 61 cucumber WRKY genes. Phylogenetic and synteny analyses were performed using related species to investigate the evolution of the cucumber WRKY genes, Sixty one WRKY genes were identified in cucumber, and insight into their classification, evolution, and expression patterns was gained in this study (**Chen *et al.*2020**)

While abiotic stress negatively affects the development and growth of plants, many studies have explored the mechanism and role of WRKY protein family members from agricultural crops, and this encapsulates that the WRKY gene family is a plant-specific transcription factor (TF) group, playing pivotal roles in many different response signalling pathways of diversifies abiotic stresses**(Li*et al.*2020)**

The transcription factor family WRKY is believed to play a key role in transcriptional reprogramming in the context of stress responses in plants. WRKY genes change their expression patterns and/or activity, contributing to the evolution of different signalling pathways and regulatory networks, through protein-protein interaction and cross-regulation and autoregulation, WRKY genes are extensively recorded, which helps us understand the complex mechanisms of signalling and transcriptional reprogramming controlled by WRKY proteins. **(Chen *et al.*2012)**

Significance of the study

The adverse effects of climate change, drought (dehydration stress), water logging (excessive watering), cold and heat (cold, frost, heat), salinity, and mineral toxicity are seen in plant development, yield, and

seed quality. In future, freshwater scarcity is expected to increase, which will inevitably increase the severity of abiotic stresses. Therefore, abiotic stress-resilient crop varieties are urgently needed to ensure food security and safety in the coming time although current knowledge about oat's stress-adaptive mechanisms is limited on a molecular level. Future findings of this study will facilitate us to understand the mechanism behind the natural adaptation of *Avena sativa* under various abiotic stresses. Furthermore, WRKY genes from *Avenasativa*could be used to design genetically improved varieties to boost agricultural productivity.

Objectives:

Avena sativa is the third major food crop growing in varied climate conditions under stress. Keeping in view the role of WRKY gene family in stress tolerance the study will be undertaken to analyse the effect of stress in respective transcription factor families and to fulfil the gap of informationfor developing stress tolerant genetically modified variety of agriculture benefits.

1) Genome wide mining for identificationand *In Silico* characterization of WRKYtranscription factor family of *Avena sativa*.
2) To study the effect of various abiotic stresses on*Avena sativa*.
3) Comparative Expression analysis of WRKY transcription factor families in control and treated *Avena sativa*.

Proposed Methodology:

1. *In Silico* database mining will be done for the identification of WRKY proteins from *Avenasativa*:

- Retrieval of entire proteome dataset, alignment of sequences.
- Construction of HMM profile.
- Identification of desired protein.
- Confirmation of the integrity of Domain, by using **HMMscan, CDD, SMART tools.**
- Identification of Physical and chemical properties of all the identified proteins of WRKY proteins.
- Prediction of Subcellular Localization of the obtained proteins.

2. Sequence analysis and Phylogenetic tree construction:

Alignment of identified multiple sequences will be carried out using **MUSCLE** with default parameters. For theConstruction of Phylogenetic tree with various desired parameters **MEGA v 7.0 software** will be used.

3. Mapping of chromosomes:

Location and positions of the predicted genes will be performed by **MapInspect v1.0 software.**

4. Visualization of gene structure and distribution of conserved motif:

Based on coding sequences, exonic and intronic positions of the identified protein will be predicted with the **GSDS server v2.0.** conserved motif will be identified with the program **MEME suite v5.1.0.** with various desired parameters.

5. Collinearity analysis with ks/ka calculation:

Homologous pairs, similarity between sequences, alignment of orthologous sequences will beperformed for the collinearity analysis to determine the synonymous (ks) and nonsynonymous (ka) substitution rates.

6. Analyses of Gene ontology (GO) along with identification of *cis-regulatory* elements.

7. Comparative modelling of protein three-dimensional structures will be performed by MODELLER.

8. Various Stress treatmentwill be provided in*Avena sativa*at different concentrations. The experimentwill be performed in triplicates:

- Abiotic stresses(Drought, Heat and Salinity).
- Phytohormones treatmentSalicylic acid (SA), Methyl Jasmonate (MeJa), Abscisic acid (ABA), Gibberellic acid (GA).

9. Comparative analysis:

- Comparative analysis of treated *Avena sativa* WRKY genes with model plants.
- Expression analysis of WRKY genes in response to Abiotic Stress.
- Expression analysis of WRKY genes in response to Hormonal Treatment.

9.Data validation, Interpretation, and conclusion.

Table 1: Various Tools and Servers

Identification oftranscription factor families through various Bioinformatics tools	**Various Tools and Servers**
1. In-silico database mining for the identification of WRKY proteins from *Avena sativa*	**Clustal Omega, HMM profile using HMMER tool v3.2**
2. Phylogenetic analysis	**MUSCLE, MEGA v7.0 software**
3. Chromosomal mapping	**MapInspect v1.0 software**
4. Gene structure and distribution of conserved motifs	**GSDS web server v2.0, MEME suite v5.1.0, TBtools**
5. Collinearity analysis and Ks/Ka calculation	**EMBOSS Needle program, PAL2NAL server**
6. GO analysis and identification of cis-regulatory elements	**Blast2GO v5 tool followed by interproscan, BLASTp.**
7. Homology Modelling and Prediction of 3rd protein structure.	**MODELLER**

References

1. Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., Pozniak, C.J., Choulet, F., Distelfeld, A., Poland, J. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science, 361.
2. Aredo, D. (1995). Famine and food security in Ethiopia. *Agricultural Economics*, *13*(1), 71-72.
3. Bai, J., Yan, W., Wang, Y., Yin, Q., Liu, J., Wight, C., & Ma, B. (2018). Screening oat genotypes for tolerance to salinity and alkalinity. *Frontiers in plant science*, *9*, 1302.
4. Bakshi, M., &Oelmüller, R. (2014). WRKY transcription factors: jack of many trades in plants. Plant Signal Behav 9: e27700.
5. Beddington, J. R., Asaduzzaman, M., Bremauntz, F. A., Clark, M. E., Guillou, M., Jahn, M. M., ... &Wakhungu, J. (2012). Achieving food security in the face of climate change: Final report from the Commission on Sustainable Agriculture and Climate Change.
6. Eulgem, T., Rushton, P. J., Robatzek, S., &Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends in plant science*, *5*(5), 199-206.
7. Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R., Dunn, M., ... & Briggs, S. (2002). A draftsequence of the rice genome (Oryza sativa L. ssp. japonica). *Science*, *296*(5565), 92-100.

8. Heuschele, D. J., Case, A., & Smith, K. P. (2019). Evaluation of fast generation cycling in oat (Avena sativa). *Cereal Research Communications*, *47*(4), 626-635.
9. Jogawat, A. (2019). Crosstalk among phytohormone signaling pathways during abiotic stress. *Molecular plant abiotic stress: biology and biotechnology*, 209-220.
10. Lamaoui, M., Jemo, M., Datla, R., &Bekkaoui, F. (2018). Heat and drought stresses in crops and approaches for their mitigation, Front. Chem. 6 (2018) 26.
11. Nolan, T., Chen, J., & Yin, Y. (2017). Cross-talk of Brassinosteroidsignaling in controlling growth and stress responses. *Biochemical Journal*, *474*(16), 2641-2661.
12. Pandey, P., Irulappan, V., Bagavathiannan, M. V., & Senthil-Kumar, M. (2017). Impact of combined abiotic and biotic stresses on plant growth and avenues for crop improvement by exploiting physio-morphological traits. *Frontiers in plant science*, *8*, 537.
13. Panta, S., Flowers, T., Lane, P., Doyle, R., Haros, G., &Shabala, S. (2014). Halophyte agriculture: success stories. *Environmental and experimental botany*, *107*, 71-83.
14. Rushton, P. J., Somssich, I. E., Ringler, P., & Shen, Q. J. (2010). WRKY transcription factors. *Trends in plant science*, *15*(5), 247-258.
15. Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., ... &Presting, G. G. (2009). The B73 maize genome: complexity, diversity, and dynamics. *science*, *326*(5956), 1112-1115.
16. Shao, H., Wang, H., & Tang, X. (2015). NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. *Frontiers in plant science*, *6*, 902.
17. Swarajyalakshmi, G., Gurumurthy, P., &Subbaiah, G. V. (2003). Soil salinity in South India: problems and solutions. *Journal of crop production*, *7*(1-2), 247-275.
18. Wang, Y. X., Liu, Z. W., Wu, Z. J., Li, H., Wang, W. L., Cui, X., & Zhuang, J. (2018). Genome-wide identification and expression analysis of GRAS family transcription factors in tea plant (Camellia sinensis). *Scientific Reports*, *8*(1), 1-19.
19. Yu, J., Hu, S., Wang, J., Wong, G. K. S., Li, S., Liu, B., ... & Yang, H. (2002). A draft sequence of the rice genome (Oryza sativa L. ssp. indica). *science*, *296*(5565), 79-92.

20. Chanwala, J., Satpati, S., Dixit, A., Parida, A., Giri, M. K., & Dey, N. (2020). Genome-wide identification and expression analysis of WRKY transcription factors in pearl millet (Pennisetum glaucum) under dehydration and salinity stress. *BMC genomics*, *21*(1), 1-16.
21. Zhang, C., Wang, D., Yang, C., Kong, N., Shi, Z., Zhao, P., ... & Chen, Q. (2017). Genome-wide identification of the potato WRKY transcription factor family. *PloS one*, *12*(7), e0181573.
22. Mao, P., Jin, X., Bao, Q., Mei, C., Zhou, Q., Min, X., & Liu, Z. (2020). WRKY transcription factors in Medicago sativa L.: Genome-wide identification and expression analysis under abiotic stress. *DNA and Cell Biology*, *39*(12), 2212-2225.
23. Muthamilarasan, M., Bonthala, V. S., Khandelwal, R., Jaishankar, J., Shweta, S., Nawaz, K., & Prasad, M. (2015). Global analysis of WRKY transcription factor superfamily in Setaria identifies potential candidates involved in abiotic stress signaling. *Frontiers in plant science*, *6*, 910.
24. Jing, Z., & Liu, Z. (2018). Genome-wide identification of WRKY transcription factors in kiwifruit (Actinidia spp.) and analysis of WRKY expression in responses to biotic and abiotic stresses. *Genes & genomics*, *40*(4), 429-446.
25. Waqas, M., Azhar, M. T., Rana, I. A., Azeem, F., Ali, M. A., Nawaz, M. A., ... & Atif, R. M. (2019). Genome-wide identification and expression analyses of WRKY transcription factor family members from chickpea (Cicer arietinum L.) reveal their role in abiotic stress-responses. *Genes & genomics*, *41*(4), 467-481.
26. Jiang, Y., Duan, Y., Yin, J., Ye, S., Zhu, J., Zhang, F., ... & Luo, K. (2014). Genome-wide identification and characterization of the Populus WRKY transcription factor family and analysis of their expression in response to biotic and abiotic stresses. *Journal of Experimental Botany*, *65*(22), 6629-6644.
27. Baillo, E. H., Hanif, M. S., Guo, Y., Zhang, Z., Xu, P., &Algam, S. A. (2020). Genome-wide Identification of WRKY transcription factor family members in sorghum (Sorghum bicolor (L.) moench). *PloS one*, *15*(8), e0236651.
28. Chen, C., Chen, X., Han, J., Lu, W., & Ren, Z. (2020). Genome-wide analysis of the WRKY gene family in the cucumber genome and transcriptome-wide identification of WRKY transcription factors that respond to biotic and abiotic stresses. *BMC plant biology*, *20*(1), 1-19.

29. Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., & Yu, D. (2012). The role of WRKY transcription factors in plant abiotic stresses. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1819*(2), 120-128.
30. Li, W., Pang, S., Lu, Z., &Jin, B. (2020). Function and mechanism of WRKY transcription factors in abiotic stress responses of plants. *Plants*, *9*(11), 1515.